

Figure 1.

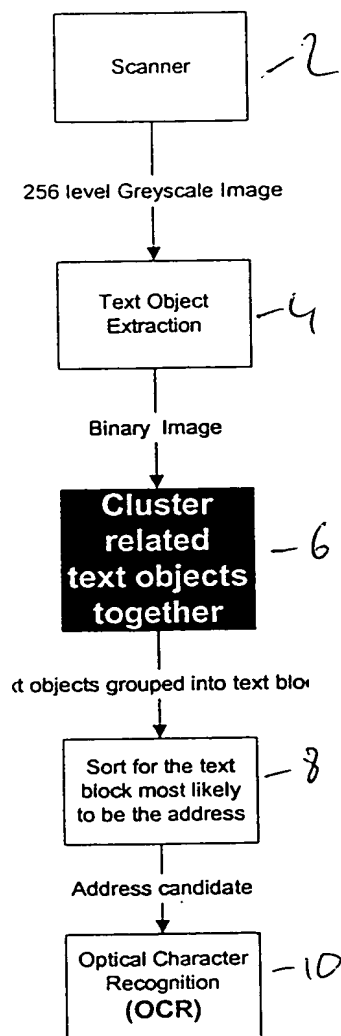


Figure 2.

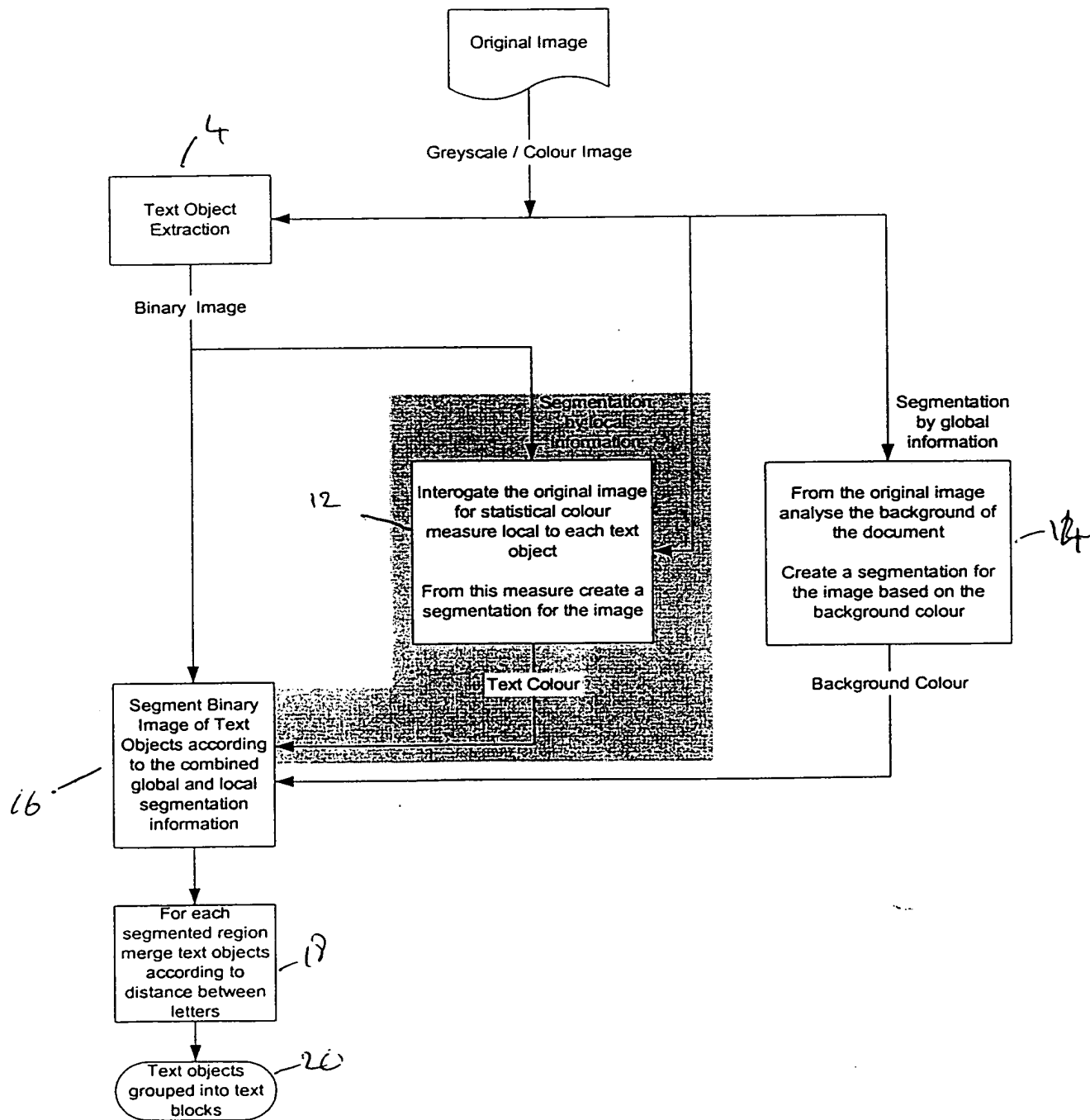
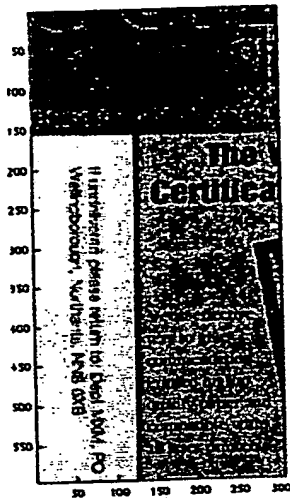


Figure 3.

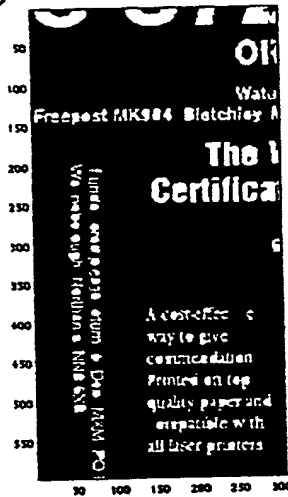
Original
Document Image

a)



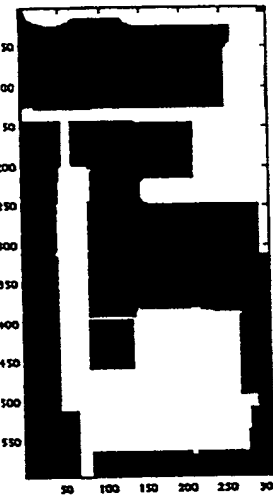
Extracted binary
text objects

b)



Clustering using simple
merging. All text
objects are clustered
together

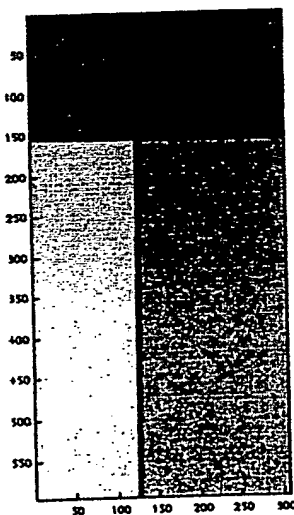
c)



3/5

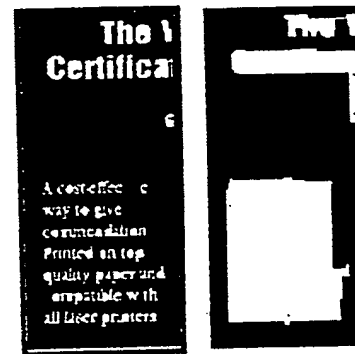
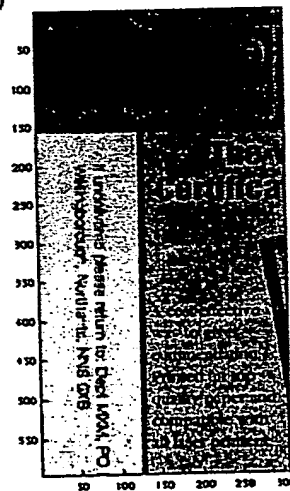
Extract
background from
original

d)



Segment according to background. For each
background region separately apply merging to the
text objects

e)



The results are clustered text objects that have
consistent background.

f)

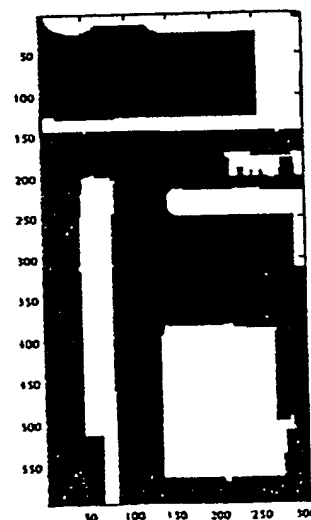


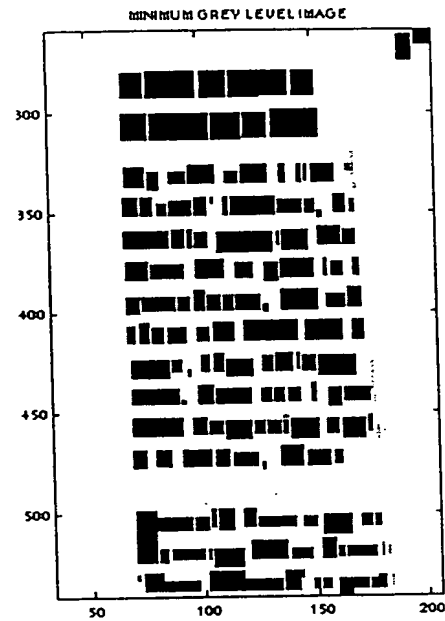
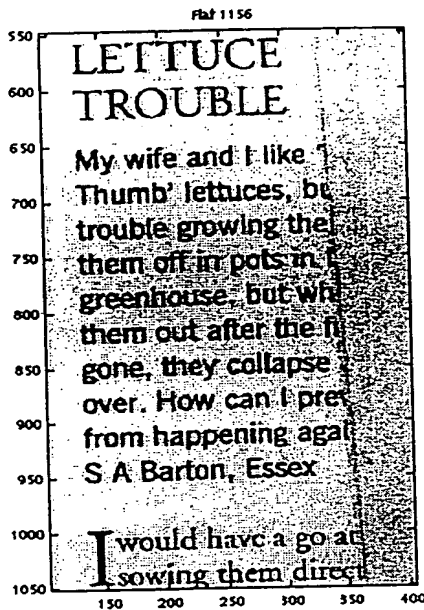
Figure 4.

4/5

For each of the binary text objects extracted the local minimum greylevel is obtained from the original document image.

The local minimum greylevel is a measure of the text colour

Original Document Image



The minimum greylevel data is used to build a segmentation of the image. Each region in segmentation is an area of the image where the text colour is consistent.

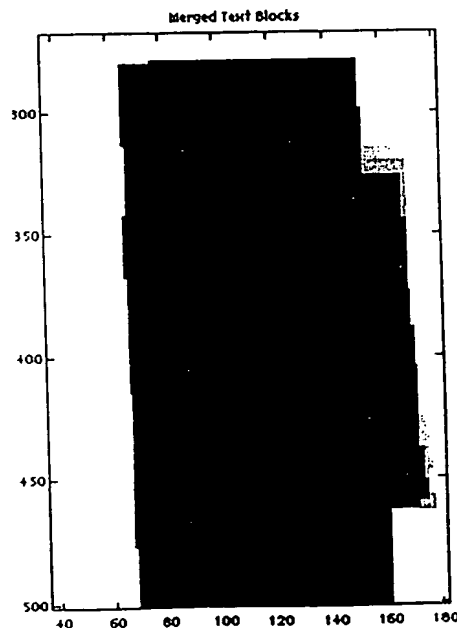


Figure 5.

5/5

